Class 19: Conditioning

Tuesday, November 19, 2019 11:38 AM

- HW 7 due Tues, Nov 26, 1pm

<u>Recall</u>: If X and Y are continuous RVs with joint PDF $f_{X,Y}$ and marginals f_X and f_Y, then the conditional PDF of

① X given Y is $f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)}$,

② Y given X is $f_{Y|X}(y|x) = \dfrac{f_{X,Y}(x,y)}{f_X(x)}$.

<u>Multiplication rule</u>:
$$f_{X,Y}(x,y) = f_{X|Y}(x|y)\, f_Y(y)$$
$$= f_{Y|X}(y|x)\, f_X(x).$$

<u>Ex 1</u>: (Filtering)

- Cop using radar gun to detect speed. Driver is going at speed $X \sim \text{Exp}(1)$,
- Given that actual speed is $X = x$, radar gun shows a speed of $Y \sim N(x, \frac{x}{2})$.
- In other words, the "error" $Y - X$ is $N(0, \frac{x}{2})$ given $X = x$.
- Signal + error = $X + (Y - X) = Y$ = observed speed
- Typical goal in this kind of problem is to study distribution of <u>X given Y</u>. (Note we are given the distr. of Y given X.)

(Note we are given the distr. of Y given X.)

Given

- $X \sim \text{Exp}(1) \rightarrow$ PDF $f_X(x) = \begin{cases} e^{-x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$
- $(Y \mid X = x) \sim N(x, \frac{x}{2})$

  $\rightarrow$ conditional PDF

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{\pi x}} e^{-\frac{(y-x)^2}{x}}$$

for $y \in \mathbb{R}$ and $x \geq 0$.

$$\boxed{\begin{array}{l} \text{PDF of } N(\mu, \sigma^2) \\ \text{is } \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \\ \text{for } y \in \mathbb{R} \end{array}}$$

Thus, from multiplication rule, joint PDF is

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) \, f_X(x)$$

$$= \begin{cases} \frac{1}{\sqrt{\pi x}} e^{-x - \frac{(y-x)^2}{x}} & \text{if } x \geq 0,\ y \in \mathbb{R} \\ 0 & \text{otherwise.} \end{cases}$$

Bayes rule:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}$$

This gives us the PDF of X given $Y = y$, which is the distribution of true speed the driver given the reading on the radar gun.

Problem: What is f_Y? (Only missing piece.)

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx = \int_0^{\infty} \frac{1}{\sqrt{\pi x}} e^{-x - \frac{(y-x)^2}{x}} dx$$

$= ???$

Conditional Expectation as a Random Variable

(a.k.a. being systematic about conditioning)

We know if (X,Y) is discrete ① or continuous ②,

then

① $\mathbb{E}[X|Y=y] = \sum_x x\, p_{X|Y}(x|y)$ where $p_{X|Y}(x|y) = P(X=x|Y=y)$

② $\mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\,dx$ where $f_{X|Y}$ is conditional PDF.

Law of total probability/expectation

① $\mathbb{E}[X] = \sum_y \underbrace{\mathbb{E}[X|Y=y]}\, p_Y(y) = \sum_y g(y)\, p_Y(y) = \mathbb{E}[g(Y)]$

② $\mathbb{E}[X] = \int_{-\infty}^{\infty} \underbrace{\mathbb{E}[X|Y=y]}\, f_Y(y)\,dy = \int_{-\infty}^{\infty} g(y) f_Y(y)\,dy = \mathbb{E}[g(Y)]$

Proof:

Recall definitions:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx, \qquad \mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx.$$

So

$$\int_{-\infty}^{\infty} \mathbb{E}[X|Y=y] f_Y(y)\,dy = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx\right) f_Y(y)\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \underbrace{f_{X|Y}(x|y)\, f_Y(y)}_{f_{X,Y}(x,y)}\,dy\,dx$$

$$= \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy}_{f_X(x)}\,dx$$

$$= \int_{-\infty}^{\infty} x f_X(x)\,dx$$

$$= \mathbb{E}[X]$$

$$= \int_{-\infty}^{\infty} x f_X(x)\,dx$$

$$= \mathbb{E}[X]$$

Definition: For each $y \in \mathbb{R}$ let $g(y) = \mathbb{E}[X \mid Y=y]$.

Plug in the random variable Y, into this function $g: \mathbb{R} \to \mathbb{R}$,

to get a new random variable $g(Y)$.

Notation: Write $\mathbb{E}[X \mid Y]$ for this RV $g(Y)$.

no "$=y$" in this notation

Note:

$$g(Y) \neq \mathbb{E}[X \mid Y=Y] = \mathbb{E}[X]$$

Tower property:

$$\mathbb{E}[\underbrace{\mathbb{E}[X \mid Y]}_{\text{best guess for } X \text{ after observing } Y}] = \underbrace{\mathbb{E}[X]}_{\text{best guess for } X}$$

best guess of the best guess...

More general versions... e.g.

$$\boxed{\mathbb{E}\left[\mathbb{E}\left[\mathbb{E}[X \mid Y, Z] \mid Y\right]\right] = \mathbb{E}[X]}$$

Ex 1 revisited:

We had $X \sim \text{Exp}(1)$ and $(Y \mid X=x) \sim N(x, \frac{x}{2})$.

Recall the mean of a $N(\mu, \sigma^2)$ random variable is μ.

Thus $\mathbb{E}[Y \mid X=x] = x$ for all $x \in \mathbb{R}$.

Conditional expectation as a random variable:

Conditional expectation as a random variable:

$$\mathbb{E}[Y|X] = X,$$

Tower property:

$$\mathbb{E}[Y] = \mathbb{E}[\overbrace{\mathbb{E}[Y|X]}] = \mathbb{E}[X] = 1.$$

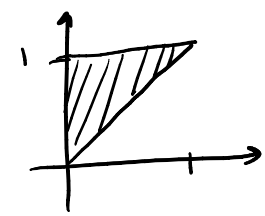
since $X \sim Exp(1)$

"Direct approach" would be

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$$

but we got stuck trying to compute f_Y before!

<u>Ex 2</u>: Suppose (X,Y) have joint PDF



$$f_{X,Y}(x,y) = \begin{cases} 8xy & \text{if } 0 \le x \le y \le 1 \\ 0 & \text{otherwise,} \end{cases}$$

we found before:

$$f_Y(y) = \begin{cases} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx = \int_0^y 8xy\, dx = 4y^3 & \text{for } 0 \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Conditional PDF of X given Y:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{8xy}{4y^3} = \frac{2x}{y^2} \quad \text{for } 0 \le x \le y \le 1.$$

<u>Calculate:</u> $\mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\, dx$

$$= \int_0^y x \cdot \frac{2x}{y^2}\, dx = \frac{1}{y^2}\left[\frac{2}{3}x^3\right]_{x=0}^{y}$$

$$= \frac{2}{3}y, \quad \text{for each } 0 \le y \le 1.$$

Conditional expectation as a RV:

$$\mathbb{E}[X|Y] = \frac{2}{3}Y.$$

Tower property:

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[\tfrac{2}{3}Y] = \tfrac{2}{3}\mathbb{E}[Y]$$

$$= \frac{2}{3}\int_{-\infty}^{\infty} y f_Y(y)\,dy$$

$$= \frac{2}{3}\int_0^1 y \cdot 4y^3\,dy = \frac{2}{3}\cdot\left[\frac{4}{5}y^5\right]_{y=0}^{1}$$

$$= \frac{8}{15}.$$

"Direct method" would be to compute f_X first then

$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx = \dots$ would give $\frac{8}{15}$ also.

This works just as easily here (no need for tower property).

Conditional Variance:

Last time we defined $Var(X|Y=y) = \mathbb{E}[X^2|Y=y] - (\mathbb{E}[X|Y=y])^2$.

Conditional Var as a RV is defined by

$$Var(X|Y) = \mathbb{E}[X^2|Y] - (\mathbb{E}[X|Y])^2.$$

$$= \mathbb{E}[(X - \mathbb{E}[X|Y])^2 \,|\, Y].$$

exercise

Useful properties of $\mathbb{E}[X|Y]$: (Linearity)

① $\mathbb{E}[X + c(Y)\,|\,Y] = \mathbb{E}[X|Y] + c(Y)$

② $\mathbb{E}[c(Y)X\,|\,Y] = c(Y)\,\mathbb{E}[X|Y]$

Recall:
$\mathbb{E}[X+c] = \mathbb{E}[X]+c$
$\mathbb{E}[cX] = c\,\mathbb{E}[X]$

$$\mathbb{E}[X+c] = \mathbb{E}[X]+c$$
$$\mathbb{E}[cX] = c\,\mathbb{E}[X]$$

② $\mathbb{E}[c(Y)X \mid Y] = c(Y)\,\mathbb{E}[X \mid Y]$

for any RV $c(Y)$ depending only on Y.

Law of total variance:

$$\underbrace{\mathrm{Var}(X)}_{\text{unconditional}} = \mathbb{E}[\underbrace{\mathrm{Var}(X \mid Y)}_{\text{conditional}}] + \mathrm{Var}(\mathbb{E}[X \mid Y])$$

Proof:

$$\mathbb{E}[\mathrm{Var}(X \mid Y)] = \mathbb{E}\left[\mathbb{E}[X^2 \mid Y] - (\mathbb{E}[X \mid Y])^2\right]$$ definition of $\mathrm{Var}(X|Y)$

$$= \mathbb{E}[\mathbb{E}[X^2 \mid Y]] - \mathbb{E}[(\mathbb{E}[X \mid Y])^2]$$ linearity of $\mathbb{E}$

$$= \mathbb{E}[X^2] - \mathbb{E}[(\mathbb{E}[X \mid Y])^2].$$ tower property

$$\mathrm{Var}(\mathbb{E}[X \mid Y]) = \mathbb{E}[(\mathbb{E}[X \mid Y])^2] - (\mathbb{E}[\mathbb{E}[X \mid Y]])^2$$ definition $\mathrm{Var}(Z) = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2$

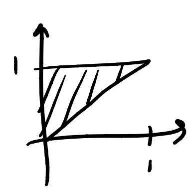
$$= \mathbb{E}[(\mathbb{E}[X \mid Y])^2] - (\mathbb{E}[X])^2$$ tower property

Adding these two, the green terms cancel out,

so $$\mathbb{E}[\mathrm{Var}(X \mid Y)] + \mathrm{Var}(\mathbb{E}[X \mid Y]) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$
$$= \mathrm{Var}(X).$$

Ex 2 revisited:

$$f_{X,Y}(x,y) = \begin{cases} 8xy & \text{if } 0 \le x \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$



Found $f_{X|Y}(x|y) = \frac{2x}{y^2}$ for $0 \le x \le y \le 1$, Found $f_{X|Y}(x|y) = \frac{2x}{y^2}$ for $0 \le x \le y \le 1$,

and $\mathbb{E}[X|Y=y] = \frac{2}{3}y$.

Also $$\mathbb{E}[X^2|Y=y] = \int_{-\infty}^{\infty} x^2 f_{X|Y}(x|y)\,dx$$

$$= \int_0^y x^2 \cdot \frac{2x}{y^2}\,dx$$

$$= \frac{1}{y^2}\left[\frac{1}{2}x^4\right]_{x=0}^{y}$$

$$= \frac{1}{2}y^2.$$

Thus $\mathbb{E}[X|Y] = \frac{2}{3}Y$, and $\mathbb{E}[X^2|Y] = \frac{1}{2}Y^2$;

so $$Var(X|Y) = \mathbb{E}[X^2|Y] - (\mathbb{E}[X|Y])^2$$

$$= \frac{1}{2}Y^2 - \frac{4}{9}Y^2$$

$$= \frac{1}{18}Y^2.$$

Law of total variance:

$$Var(X) = \mathbb{E}[Var(X|Y)] + Var(\mathbb{E}[X|Y])$$

$$= \mathbb{E}\left[\frac{1}{18}Y^2\right] + Var\left(\frac{2}{3}Y\right)$$

$$= \frac{1}{18}\mathbb{E}[Y^2] + \frac{4}{9}Var(Y)$$

$$= \frac{1}{18}\mathbb{E}[Y^2] + \frac{4}{9}\left(\mathbb{E}[Y^2] - (\mathbb{E}[Y])^2\right)$$

$$= \left(\frac{1}{18} + \frac{4}{9}\right)\mathbb{E}[Y^2] - \frac{4}{9}(\mathbb{E}[Y])^2.$$

Then find f_Y and compute

$$\mathbb{E}[Y^2] = \int_{-\infty}^{\infty} y^2 f_Y(y)\,dy \quad \text{and} \quad \mathbb{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y)\,dy$$

[illegible] $Var(X)$

$\mathbb{E}[Y] = \int_{-\infty}$ [illegible]

and plug this in to get $Var(X)$.

In this case, the "direct method" is probably easier,

to find f_X and use it to compute $Var(X)$ directly.

(Both give same answer.)